BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 9: e180327
doi: 10.3897/biss.9.180327

OPEN ACCESS

Conference Abstract

Rescuing the Past to Prepare for the Future: Environmental Data Rescue as a Key Activity of Open Science

Diane S Srivastava[‡], David AGA Hunt[§], Sandra A Binning[|], Sandra Michelle Emry[‡], Ellen K Bledsoe[¶], Jessica E Reemeyer[#], Jason Pither[¤]

‡ University of British Columbia, Vancouver, Canada
§ Canadian Institute of Ecology and Evolution - Living Data Project, Montreal, Canada
| University of Montreal, Montreal, Canada
¶ University of Arizona, Tucson, United States of America
# University of Regina, Regina, Canada
¤ University of British Columbia, Kelowna, Canada

Corresponding author: Diane S Srivastava (srivast@zoology.ubc.ca)

Received: 25 Nov 2025 | Published: 09 Dec 2025

Citation: Srivastava D, Hunt D, Binning S, Emry S, Bledsoe E, Reemeyer J, Pither J (2025) Rescuing the Past to Prepare for the Future: Environmental Data Rescue as a Key Activity of Open Science. Biodiversity Information Science and Standards 9: e180327. https://doi.org/10.3897/biss.9.180327

Abstract

The need for data rescue. Data are dying all around us: when they are stored in inaccessible or fragile repositories, in proprietary formats, on defunct media or without complete metadata. In environmental science, such data loss has high societal and scientific costs. For example, failure to archive ecological data focused on the Exxon Valdez oil spill represents an estimated loss of >100 million USD (Bledsoe et al. 2022). Loss of environmental data can also represent erasure of historic baselines that can never be recollected.

The Living Data Project (LDP). This award-winning program helps scientists and organizations archive valuable "legacy" datasets, making them permanently open and accessible (Bledsoe et al. 2022). We pair data custodians with graduate students who we train in reproducible data management and provide mentoring by postdoctoral researchers and faculty and assistance by undergraduates. By training Canada's next generation of scientists, we aim to ensure that data are never lost again. Over the last five

© Srivastava D et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

years, we have trained 305 students, across 93 internships, rescuing > 2000 years of data. LDP student working groups then analyze archived data to investigate pressing biodiversity questions and develop open access teaching resources.

Benefits of data rescue. LDP-rescued datasets, many stretching back into the last century, provide missing historical baselines that enable the detection of environmental change – such as change in lake ice formation and melt over the last 150 years. Rescue of historical datasets, like surveys of urban birds or plant communities collected five decades ago, sets the stage for contemporary re-surveys. Combining rescued data with contemporary data can also yield discoveries such as legacy effects of pesticides on stream health (Sugden et al. 2025). LDP rescue of classic studies in ecological and evolutionary theory (Darwin's Finches, Serengeti ungulates, and freshwater sticklebacks) enables new analyses that were computationally impossible at the time. LDP projects help environmental organizations archive data over broad spatial and temporal scales, critical for metrics such as the Watershed Reports or Living Planet Index. LDP also assists non-professionals to archive data; e.g., helping communities archive water quality data on DataStream enables easy comparison to national standards. Data rescue can knit together disparate datasets in relational databases, such as LDP compilation of fish, invertebrate, phytoplankton, and limnology data from the Canadian government's Turkey Lakes Watershed research. The benefits of data rescue will pay incalculable dividends for decades to come. However, to ensure successful data rescue, we recommend the following:

Rules for success:

1. Prevent scope creep. When non-essential data is included in the initial scope of the project or the scope broadens during the rescue process, concluding the project is challenging. It is better to successfully archive a subset of the data than fail to archive any.
2. Involve data collectors and custodians throughout. They provide important contextual knowledge, help decipher obscure data codes and can suggest constraints to values, facilitating validation. Formalizing their knowledge in the metadata ensures future usability.
3. Continually champion projects until the dataset DOI is minted. It is easy for projects to stall in the final stages. Delaying archiving to add more years or variables or to publish one more paper increases the risk of no data being archived. Encourage the use of versioning or embargo periods, available for many repositories, to deposit promptly the core data while enabling later additions and public release.
4. Manage adaptively. Problems like contradictory data, inconsistent use of terms, and heterogeneous collection methods may only be apparent after working with the data. At the LDP, we use high-level checks of progress throughout each rescue project to adaptively manage the scope and personnel needed, informed by our experience in a wide variety of projects.

Summary. Environmental data rescue has many benefits, including establishing baselines and change in the environment, enabling new analyses and composite indices, connecting disparate data, facilitating community science, and training the next generation of researchers in open science. However, data rescue projects requires continued and careful management to ensure success, which we defined as the deposition of data in open, accessible and permanent repositories.

Keywords

data loss, historical baselines, training, data archiving, research data management, open data

Presenting author

Diane S Srivastava

Presented at

Living Data 2025

Funding program

Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Training Experience program (CREATE)

Grant title

NSERC CREATE in ecology, evolution and environmental science: The Living Data Project (2020-543234)

Author contributions

DSS: writing (original draft); all other authors: writing (review & editing).

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Bledsoe EK, Burant JB, Higino GT, Roche DG, Binning SA, Finlay K, Pither J, Pollock LS, Sunday JM, Srivastava DS (2022) Data rescue: saving environmental data from extinction. Proceedings of the Royal Society B 289 (1979): 20220938. https://doi.org/10.1098/rspb.2022.0938
- Sugden S, White AB, Lento J, Kurek J, Dimitrovas I, Emry S, Hua X, Ijzerman MM, Kidd KA, Morrow KL, et al. (2025) Legacy effects of four decades of insecticide applications on contemporary riverine benthic macroinvertebrates. Environmental Pollution 376: 126397. https://doi.org/10.1016/j.envpol.2025.126397